

中華民國經濟部智慧財產局

INTELLECTUAL PROPERTY OFFICE  
MINISTRY OF ECONOMIC AFFAIRS  
REPUBLIC OF CHINA

茲證明所附文件，係本局存檔中原申請案的副本，正確無訛，  
其申請資料如下：

This is to certify that annexed is a true copy from the records of this  
office of the application as originally filed which is identified hereunder:

申 請 日：西元 2003 年 04 月 14 日  
Application Date

申 請 案 號：092108577  
Application No.

申 請 人：財團法人工業技術研究院  
Applicant(s)

局 長  
Director General

蔡 練 生

發文日期：西元 2003 年 5 月 5 日  
Issue Date

發文字號：09220439230  
Serial No.

# 發明專利說明書

(填寫本書件時請先行詳閱申請書後之申請須知，作※記號部分請勿填寫)

※ 申請案號：\_\_\_\_\_ ※IPC分類：\_\_\_\_\_

※ 申請日期：\_\_\_\_\_

## 壹、發明名稱

(中文) 詞語驗證方法及系統

(英文) METHOD AND SYSTEM FOR UTTERANCE VERIFICATION

## 貳、發明人(共2人)

發明人1 (如發明人超過一人，請填說明書發明人續頁)

姓名：(中文) 張森嘉

(英文) Sen-Chia Chang

住居所地址：(中文) 新竹市明湖路1200巷86弄31號

(英文) 31, Alley 86, Lane 1200, Ming-Hu Road, Hsin-Chu, Taiwan R.O.C.

國籍：(中文) 中華民國 (英文) R.O.C.

## 參、申請人(共1人)

申請人1 (如發明人超過一人，請填說明書申請人續頁)

姓名或名稱：(中文) 財團法人工業技術研究院

(英文) Industrial Technology Research Institute

住居所或營業所地址：(中文) 新竹縣竹東鎮中興路四段一九五號

(英文) No. 195, Sec. 4, Chung Hsing Rd., Chutung, Hsinchu

國籍：(中文) 中華民國 (英文) R.O.C.

代表人：(中文) 翁政義

(英文) Cheng-I Weng

繢發明人或申請人續頁 (發明人或申請人欄位不敷使用時，請註記並使用續頁)

發明人 2

姓名：(中文) 簡世杰

(英文) Shih-Chih Chieh

住居所地址：(中文) 台中縣霧峰鄉坑口村復興六街一巷 12 號

(英文) No. 12, Lane 1, Fushing 6 St., Wufeng, Taichung, Taiwan R. O. C.

國籍：(中文) 中華民國

(英文) R.O.C.

發明人 3

姓名：(中文)

(英文)

住居所地址：(中文)

(英文)

國籍：(中文)

(英文)

發明人 4

姓名：(中文)

(英文)

住居所地址：(中文)

(英文)

國籍：(中文)

(英文)

發明人 5

姓名：(中文)

(英文)

住居所地址：(中文)

(英文)

國籍：(中文)

(英文)

發明人 6

姓名：(中文)

(英文)

住居所地址：(中文)

(英文)

國籍：(中文)

(英文)

## 肆、中文發明摘要

本發明係有關一種詞語驗證方法及系統，其首先抽取  
出語音信號中的特徵參數向量序列，再經過語音辨識後可  
取得至少一候選詞，依照候選詞的詞彙內容所對應的驗證  
單元將該語音信號切割為對應於驗證單元的語音音段，並  
求出這些語音音段的驗證用特徵參數向量序列，之後依序  
使用這些驗證用特徵參數向量序列進行驗證，以產生驗證  
分數；這個驗證方法係使用語音音段所相對應之驗證單元  
的類神經網路來計算驗證分數，而這個類神經網路係為一  
多層感知，驗證分數係使用驗證用特徵參數向量序列對此  
多層感知進行前饋動作求得；在合併所有語音音段之驗證  
分數以取得詞語驗證分數後，即可根據一預先定義好的門  
檻值來決定接受或拒絕該候選詞。

## 伍、英文發明摘要

A method and system for utterance verification is disclosed. It can extract a sequence of feature vectors from speech signal. At least one candidate string is obtained after the process of speech recognition by using the sequence of feature vectors. Then, the speech signal is segmented into speech segments according to the verification-unit-specified structure of candidate string, and that makes each speech segment corresponding to a verification-unit. After calculating the sequences of verification-specific feature vectors for these speech segments, sequentially use these sequences to generate the verification scores of speech segments in verification process. This invention uses neural networks for the calculations of verification scores. Where each neural network is a Multi-Layer Perceptron (MLP) and was developed for each verification-unit. Verification score is obtained through using the feed-forward process of MLP. Finally, utterance verification score is obtained by combining all of the verification scores of speech segments and is used to compare with a pre-defined threshold for the decision of acceptance or rejection of the candidate string.

陸、(一)、本案指定代表圖爲：圖 4

(二)、本代表圖之元件代表符號簡單說明：

流程圖無元件代表符號說明

柒、本案若有化學式時，請揭示最能顯示發明特徵的化學式：

無



## 捌、聲明事項

本案係符合專利法第二十條第一項第一款但書或第二款但書規定之期間，其日期為：\_\_\_\_\_

本案已向下列國家（地區）申請專利，申請日期及案號資料如下：

【格式請依：申請國家（地區）；申請日期；申請案號 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

主張專利法第二十四條第一項優先權：

【格式請依：受理國家（地區）；日期；案號 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_
5. \_\_\_\_\_
6. \_\_\_\_\_
7. \_\_\_\_\_
8. \_\_\_\_\_
9. \_\_\_\_\_
10. \_\_\_\_\_

主張專利法第二十五條之一第一項優先權：

【格式請依：申請日；申請案號 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

主張專利法第二十六條微生物

國內微生物 【格式請依：寄存機構；日期；號碼 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

國外微生物 【格式請依：寄存國名；機構；日期；號碼 順序註記】

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

熟習該項技術者易於獲得，不須寄存。

## 玖、發明說明

(發明說明應敘明：發明所屬之技術領域、先前技術、內容、實施方式及圖式簡單說明)

### 一、發明所屬之技術領域

本發明係關於一種詞語驗證方法及系統，尤指一種具有高可靠度之詞語驗證方法及系統，其可適用於噪音環境  
5 下。

### 二、先前技術

按，詞語驗證(utterance verification)技術係用於驗證一段語音信號經過語音辨識後所得到的候選詞的正確性，  
10 倘使系統根據驗證無誤的候選詞加以作動。例如應用於語音撥號系統中，當系統接收到包含電話號碼的語音信號時，將會辨識並驗證語音信號中所包含的數字，藉以根據驗證無誤後之數字來進行撥接。

習知之詞語驗證技術中，最被廣泛使用的方法主要包括兩大類，分別為使用解譯(decoding)技術來進行詞語驗證、以及以假說檢定(hypothesis testing)為基礎的詞語驗證方法。

請參閱第1圖關於習知解譯技術應用於詞語驗證之方塊圖，顯示自輸入之語音信號51中辨識出英文單字(word)  
20 『Hi』，包含有『h』及『ai』兩個部分。由於在解譯技術中通常都是將『Hi』視為一個單元來考慮，因此這種解譯技術所使用的參數52係以『Hi』為單元來計算，並且常使用一個以上的參數來進行驗證分數的計算；包括『Hi』的聲學分數(acoustic score)521、語言模型分數(language model score)522、以及詞語的辨識排名資訊(N-Best

information)523...等。之後，透過解譯器(decoder)53將上述分數結合、以計算出『Hi』的驗證分數(verification score)54，最後再藉由比較驗證分數54與預設之門檻值來決定接受或拒絕該辨識結果。其中，上述之解譯技術可使用例如線性驗證分析(Linear Discriminative Analysis, LDA)、決策樹分析(decision tree analysis)、或類神經網路(neural networks)等方法來進行。然而，由於這類方法需使用到各種不同的參數，且參數求取的複雜度較高，因此，對系統資源的要求也較高。

第2圖則為習知以假說檢定為基礎來進行詞語驗證之方塊圖，語音信號61被辨識且得到英文單字『Hi』。根據假說檢定的方法，必須先把語音信號依據辨識出之詞彙內容切割成對應於辨識單元(recognition unit)的語音音段(speech segment)後再加以驗證。例如第1圖所示之單字『Hi』，即依據『Hi』的詞彙內容將該語音信號切割為對應於以次詞(subword)為辨識單元的語音音段，即『h』及『ai』音段。接著，分別使用次詞『h』及『ai』驗證模型(verification model)<sup>621,623</sup>與反向模型(anti-model)<sup>622,624</sup>，來對次詞『h』及『ai』所對應的語音音段進行檢定，以根據兩兩模型的機率比值來計算出『h』的檢測分數<sup>631</sup>與『ai』的檢測分數<sup>632</sup>。最後再藉由結合檢測分數<sup>631,632</sup>以形成『Hi』的驗證分數<sup>64</sup>，並與預設之門檻值相比較來決定接受或拒絕該辨識結果。然而，上述以假說檢定為基礎之驗證方法必須分別為每個辨識單元建

立驗證模型與反向模型；且每個語音音段的驗證都必須經過兩次檢定，將大幅增加系統負荷。

此外，上述解譯方法及假說檢定方法多半實施於無噪音干擾的環境，因此一旦應用在噪音環境中時，將嚴重影響詞語驗證的正確率，並降低辨識結果的可靠性，繼而去語言辨識系統的親和力。以上均非十分理想。

發明人爰因於此，本於積極發明之精神，亟思一種可以解決上述問題之「詞語驗證方法及系統」，幾經研究實驗終至完成此項嘉惠世人之發明。

10

### 三、發明內容

本發明之主要目的係在提供一種詞語驗證方法及系統，使其具有抗噪能力以適用於噪音環境中，俾能增進語言辨識系統在噪音環境下的可靠度以及系統的親和性。

15

本發明之另一目的係在提供一種詞語驗證方法及系統，其對於每一驗證單元僅需建立一個類神經網路，俾能降低系統負荷，並提高詞語驗證效能。

20

本發明之再一目的係在提供一種詞語驗證方法及系統，俾能輕易轉用於各種不同應用領域之語言辨識系統上，以提高這種驗證方法的可攜性。

依據本發明之一特色，於所提出之詞語驗證方法中，首先係由一輸入語音信號中抽取出一特徵參數向量序列，序列中每一特徵參數向量皆具有複數個維度(dimensional)的特徵值；其次，將前述特徵參數向量序列輸入至一語音

辨識器中，以辨識出至少一候選詞；之後，根據該候選詞的詞彙內容所對應的驗證單元以將該語音信號切割為對應於驗證單元的語音音段；再依序使用該等驗證單元之正規化參數來對該等語音音段之特徵參數向量序列進行正規化，以取得語音音段之驗證用特徵參數向量序列；接著，依照各語音音段所相對應之驗證單元的類神經網路進行驗證分數的計算；在完成驗證分數的計算後，合併這些語音音段的驗證分數以取得一詞語驗證分數；最後將此詞語驗證分數與一預設之門檻值比較，當詞語驗證分數大於門檻值時，接受此候選詞，反之則拒絕此候選詞。

依據本發明之另一特色，係提出一種詞語驗證系統，其包括有一特徵參數抽取模組、一語音辨識模組、一語音切割模組、一驗證用特徵參數製作模組、一驗證分數計算模組、一驗證分數合併模組、以及一決策模組，該特徵參數抽取模組由一輸入語音信號中抽取出一特徵參數向量序列；該語音辨識模組係自該等特徵參數向量序列中辨識出至少一候選詞；該語音切割模組係依照該等候選詞的詞彙內容所對應的驗證單元以將該輸入語音信號切割為對應於驗證單元的語音音段；該驗證用特徵參數製作模組係根據該等語音音段之特徵參數向量序列以產生驗證用特徵參數向量序列；該驗證分數計算模組係使用該等語音音段所相對應之驗證單元的類神經網路，根據該等語音音段之驗證用特徵參數向量序列來計算出該等語音音段之驗證分數；該驗證分數合併模組係合併所有語音音段之驗證分數，以

取得一詞語驗證分數；該決策模組比較該詞語驗證分數與一預設之門檻值，當該詞語驗證分數大於該門檻值時，接受該候選詞，反之則拒絕該候選詞。

由於本發明構造新穎，能提供產業上利用，且確有增進功效，故依法申請發明專利。

#### 四、實施方式

有關本發明之詞語驗證方法及系統之較佳實施例，請先參閱第3圖之功能方塊圖，其顯示本實施例係使用一電腦系統20以對所輸入之語音信號10進行詞語驗證，其中，電腦系統20包括有一特徵參數抽取模組21、一語音辨識模組22、一聲學模型資料庫28、一存有至少一個詞彙之詞彙資料庫29、一語音切割模組23、一驗證用特徵參數製作模組24、一正規化參數資料庫11、一驗證分數合併模組26、一決策模組27、以及一驗證分數計算模組25，並使用複數個類神經網路121、122、123來進行驗證分數計算。其中，本實施例之類神經網路係為一多層感知 (Multi-Layer Perceptron)。

請一併參閱第4圖之流程圖，當電腦系統20接收到一由外界所輸入之語音信號10時，便透過特徵參數抽取模組21將語音信號10視窗化 (windowing) 為複數個長度相同的音框 (frame)，以抽取出一特徵參數向量序列211(步驟S401)，俾利於進行語音辨識與詞語驗證，在序列211中，每一特徵參數向量係包括有複數個維度的特徵值。於本實

施例中，用以求取特徵參數向量序列211所使用的音框長度為160點(20ms)，音框重疊長度為80點(10ms)，並以漢明視窗(Hamming Window)做平滑化處理；且每一個特徵參數向量係具有二十六個維度的特徵值，分別為十二維的梅氏倒頻譜係數(Mel-cepstral coefficient)、十二維的差分(delta)倒頻譜係數、與對數能量的一次及二次差分(delta-log-energy及delta-delta-log-energy)；此外，本實施例更採用倒頻譜均值正規化(cepstral mean normalization)來補償通道效應。

之後將所抽取出之特徵參數向量序列211輸入至語音辨識模組22，以辨識出至少一候選詞221(步驟S402)。該語音辨識模組22係使用一以隱藏式馬可夫模型(Hidden Markov Model, HMM)為基礎的聲學模型資料庫28及一存有至少一個詞彙的詞彙資料庫29來對前述之特徵參數向量序列211進行辨識動作，以輸出至少一候選詞。其後，透過語音切割模組23依據候選詞221的詞彙內容所對應的驗證單元，將語音信號10切割為對應於驗證單元的語音音段231,232,233(步驟S403)。其中，若候選詞221為中文，則切割出來的語音音段231,232,233為對應到次音節驗證單元的次音節音段(subsyllable segment)，例如當候選詞221為『書』時，根據『書』的詞彙內容所切割出來的語音音段為對應到次音節驗證單元『戶』及『ㄨ』等兩個次音節音段；若候選詞221為英文字，則切割出來的語音音段231,232,233為對應到次詞驗證單元的次詞音段(subword

segment)，例如『sky』將會根據『sky』的詞彙內容而切割出『s』、『k』、及『ai』等對應到次詞驗證單元的次詞音段。

再來分別將每一語音音段231,232,233之特徵參數向量序列傳送至驗證用特徵參數製作模組24，進行驗證用特徵參數向量序列的製作。其中，驗證用特徵參數向量序列的製作係以每一語音音段所對應的驗證單元的正規化參數來對該語音音段之特徵參數向量序列進行正規化後求得。其中，驗證單元的正規化參數係為一組特徵參數向量的平均值及標準差，該正規化參數係預先統計訓練語料中對應到該驗證單元的語音音段的特徵參數向量序列而得到，並儲存於正規化參數資料庫11中。以語音音段231為例，其驗證用特徵參數向量序列241係自正規化參數資料庫11中取出該語音音段所對應之驗證單元的正規化參數來對該語音音段231之特徵參數向量序列進行正規化而求得(步驟S404)，其中241的每一驗證用特徵參數向量亦包括有複數個維度的特徵值。同理，可依序製作出語音音段232及233的驗證用特徵參數向量序列242及243。

接著，依序將前述驗證用特徵參數向量序列241,242,243輸入至驗證分數計算模組25，依照各語音音段231,232,233所相對應之驗證單元的類神經網路121,122,123進行驗證分數的計算。該些類神經網路121,122,123係為多層感知，驗證分數251,252,253的計算係使用類神經網路121,122,123對驗證用特徵參數向量序列

241,242,243 進行前饋 (feed-forward) 動作而求得 (步驟 S405)。為清楚描述驗證分數的計算，以下舉本實施例使用  
類神經網路 121 計算驗證分數 251 之過程來詳加描述。

請一併參閱第 5 圖本實施例類神經網路 121 之示意圖，其係為一多層感知，包括有一輸入層 31、一隱藏層 32、以及一輸出層 33，且每一驗證單元係對應至一類神經網路。其中，輸入層 31 之輸入神經元 (input neuron) 311 用以依序接收驗證用特徵參數向量序列 241 裡每一個驗證用特徵參數向量的特徵值  $D_i$ ，並將之輸出至隱藏層 32，隱藏層 32 之隱藏神經元 (hidden neuron) 321 用以接收輸入層 31 裡所有的輸入神經元的輸出，並將結果輸出至輸出層 33，輸出層 33 僅使用一輸出神經元 (output neuron) 331 來接收隱藏層 32 裡所有的隱藏神經元的輸出，並輸出該驗證用特徵參數向量的驗證結果。因此，透過上述的前饋動作，即可求出驗證用特徵參數向量序列 241 裡每個向量的驗證結果。驗證分數 251 即是取這些驗證結果的平均值而得到。

至於類神經網路 121 中各個神經元的輸出，除了輸入神經元是直接輸出驗證用特徵參數向量的特徵值之外，其他神經元的輸出（隱藏神經元及輸出神經元）則是使用下列公式來計算：

$$out_j = \frac{1}{1 + \exp(-\sum_i w_{ji} out_i + b_j)},$$

其中， $out_j$  為本層（隱藏層或輸出層）中第  $j$  個神經元的輸出值， $out_i$  為前一層（輸入層或隱藏層）中第  $i$  個神經

元的輸出值， $w_{ji}$ 為前一層第*i*個神經元至本層第*j*個神經元的權重 (weight)， $b_j$ 為本層第*j*個神經元的偏移量(bias)。

在完成驗證分數251,252,253的計算後，接著，透過驗證分數合併模組26來計算出這些驗證分數的平均值，以取得候選詞221之詞語驗證分數261(步驟406)。最後決策模組5 27使用一預設之門檻值(threshold)與詞語驗證分數261進行比較(步驟407)，當詞語驗證分數261大於門檻值時，電腦系統20接受此候選詞221(步驟S408)，反之則拒絕(步驟S409)。

10 而為了使本實施例所使用的多層感知能夠藉由使用前述驗證用特徵參數向量序列來分辨所屬的語音音段是否為對應的驗證單元，用於訓練某個驗證單元之多層感知的訓練資料係包含了對應到該驗證單元的語音音段及不是對應到該驗證單元的語音音段，以使多層感知能夠學習到其15 間的差異，進而分辨之。在訓練上，必須先為這些驗證用特徵參數向量序列所屬的語音音段設定相對應的目標值，之後，再透過錯誤回傳(error back-propagation)學習演算法來進行訓練，藉由縮小實際輸出之驗證分數與目標值間的均方根誤差(mean square error)來調整多層感知裡的權重20 和偏移量。例如，訓練一個驗證單元為『ㄨ』的多層感知時，訓練所使用的語音音段包括了對應到驗證單元為『ㄨ』的語音音段和不是對應到驗證單元『ㄨ』的語音音段。對於對應到驗證單元為『ㄨ』的語音音段，其目標值為1，反之，其目標值為0。經由這樣的設定及使用上述的訓練方法

進行反覆訓練，將使得多層感知實際輸出的驗證分數逼近目標值，而使該多層感知能夠藉由使用驗證用特徵參數向量序列來分辨所屬的語音音段是否為對應的驗證單元。

而為了使本實施例之詞語驗證方法及其系統能夠適用於噪音環境，上述訓練多層感知所使用的訓練資料係為受到噪音干擾且具各種不同訊噪比(Signal Noise Ration, SNR)的語音音段，例如使用受車內噪音(In-Car noise)干擾後具訊噪比為9dB、3dB、0dB、-3dB、或-9dB的語音音段，以使多層感知能夠學習受到不同程度的車內噪音所干擾之語音特性，以期在實際應用時，亦能夠分辨出受到噪音干擾的語音音段是否為對應的驗證單元。

由以上說明可知，本發明以次音節或次詞作為驗證單元，因此能夠被廣泛且容易地應用到以次音節或次詞為辨識單元的語音辨識系統上，且每個驗證單元僅需建立一個多層感知，故語音音段是否對應到正確的驗證單元只要使用其所屬的驗證用特徵參數向量序列對其相對應的驗證單元的多層感知進行前饋動作即可完成，將可大幅降低電腦系統的負荷。此外，本發明係使用受到噪音干擾的語音音段來訓練多層感知，因此能增進語音辨識系統在噪音環境下的可靠度與親和性，提高了使用便利性。

綜上所陳，本發明無論就目的、手段及功效，在在均顯示其迥異於習知技術之特徵，實為一極具實用價值之發明，應符合專利要件，懇請 貴審查委員明察，早日賜准專利，俾嘉惠社會，實感德便。惟應注意的是，上述諸多

實施例僅係為了便於說明而舉例而已，本發明所主張之權利範圍自應以申請專利範圍所述為準，而非僅限於上述實施例。

## 5 五、圖式簡單說明

第1圖係習知使用解譯技術進行詞語驗證之功能方塊圖。  
第2圖係習知使用假說檢定進行詞語驗證之功能方塊圖。  
第3圖係本發明實施例詞語驗證系統之功能方塊圖。  
第4圖係本發明實施例詞語驗證方法之流程圖。  
10 第5圖係本發明實施例類神經網路之示意圖。

## 六、圖號說明

語音信號 10	正規化參數資料庫 11	
類神經網路 121,122,123	電腦系統 20	
15 特徵參數抽取模組 21	特徵參數向量序列 211	
語音辨識模組 22	候選詞 221	語音切割模組 23
對應於驗證單元的語音音段 231,232,233		
驗證用特徵參數製作模組 24		
驗證用特徵參數向量序列 241,242,243		
20 驗證分數計算模組 25		
語音音段之驗證分數 251,252,253		
驗證分數合併模組 26	詞語驗證分數 261	
決策模組 27	聲學模型資料庫 28	
詞彙資料庫 29	輸入層 31	

輸入神經元 311  
隱藏神經元 321  
輸出神經元 331  
參數 52  
5 語言模型分數 522  
解譯器 53  
語音信號 61  
反向模型 622,624  
驗證分數 64  
隱藏層 32  
輸出層 33  
語音信號 51  
聲學分數 521  
詞語辨識排名資訊 523  
驗證分數 54  
驗證模型 621,623  
檢測分數 631,632

## 拾、申請專利範圍

1. 一種詞語驗證方法，包含下列步驟：

(A) 由一輸入語音信號中抽取出一特徵參數向量序列；

5 (B) 將該特徵參數向量序列輸入至一語音辨識器中，並辨識出至少一候選詞；

(C) 依照候選詞之詞彙內容所對應的驗證單元將該輸入語音信號切割為對應於驗證單元的語音音段；

10 (D) 使用該等語音音段所對應之驗證單元之正規化參數對該等語音音段的特徵參數向量序列進行正規化，以產生驗證用特徵參數向量序列；

(E) 使用該等語音音段所相對應之驗證單元的類神經網路，根據該等語音音段之驗證用特徵參數向量序列來計算出該等語音音段之驗證分數；

15 (F) 合併所有語音音段之驗證分數，以取得一詞語驗證分數；以及

(G) 將該詞語驗證分數與一預設之門檻值比較，當該詞語驗證分數大於該門檻值時，接受該候選詞，反之則拒絕該候選詞。

20 2. 如申請專利範圍第1項所述之方法，其中，於步驟(D)中，該驗證單元之正規化參數係為一組特徵參數向量的平均值及標準差，其係預先統計訓練語料中對應到該驗證單元的語音音段的特徵參數向量序列而得到。

3.如申請專利範圍第1項所述之方法，其中，於步驟(E)中，每一驗證單元係對應至一類神經網路，且該類神經網路係為一多層感知。

4.如申請專利範圍第3項所述之方法，其中，該多層感知係使用該語音音段之驗證用特徵參數向量序列作為輸入，以一前饋動作求出序列中每一驗證用特徵參數向量的驗證結果，並取該等驗證結果的平均值為該語音音段之驗證分數。

5.如申請專利範圍第3項所述之方法，其中，該多層感知係使用一錯誤回傳學習演算法來縮小由該多層感知所輸出之驗證分數與一目標值間的均方根誤差，來訓練該多層感知。

6.如申請專利範圍第5項所述之方法，訓練對應到一個驗證單元的多層感知係藉由輸入對應到該驗證單元的語音音段之驗證用特徵參數向量序列及非對應到該驗證單元的語音音段之驗證用特徵參數向量序列來訓練該多層感知，其中，訓練所使用的語音音段若對應到該驗證單元，則目標值為1，反之，則目標值為0。

7.如申請專利範圍第1項所述之方法，其中，步驟(F)係取所有語音音段之驗證分數的平均值為該詞語驗證分數。

8.如申請專利範圍第1項所述之方法，其中，該輸入語音信號係為一受噪音干擾之語音信號。

9. 如申請專利範圍第6項所述之方法，其中，該些為訓練所使用的語音音段係為受噪音干擾之語音音段。

10. 一種詞語驗證系統，包括：

一特徵參數抽取模組，用以由一輸入語音信號中抽出  
5 出一特徵參數向量序列；

一語音辨識模組，係自該特徵參數向量序列中辨識出  
至少一候選詞；

一語音切割模組，係依照候選詞內容所對應的驗證單元將該輸入語音信號切割為對應於驗證單元的語音音段；

10 一驗證用特徵參數製作模組，係根據該等語音音段所對應之驗證單元之正規化參數對該等語音音段之特徵參數向量序列進行正規化，以產生驗證用特徵參數向量序列；

一驗證分數計算模組，係使用該等語音音段所相對應之驗證單元的類神經網路，根據該等語音音段之驗證用特  
15 徵參數向量序列來計算出該等語音音段之驗證分數；

一驗證分數合併模組，係合併所有語音音段之驗證分數，以取得一詞語驗證分數；以及

一決策模組，用以比較該詞語驗證分數與一預設之門檻值，當該詞語驗證分數大於該門檻值時，接受該候選詞，  
20 反之則拒絕該候選詞。

11. 如申請專利範圍第10項所述之系統，其中，該驗證用特徵參數製作模組所使用之驗證單元之正規化參數係為一組特徵參數向量的平均值及標準差，且係預先統計訓

練語料中對應到該驗證單元的語音音段的特徵參數向量序列而得到。

12. 如申請專利範圍第10項所述之系統，其中，每一驗證單元係對應至一類神經網路，且該類神經網路係為一  
5 多層感知。

13. 如申請專利範圍第12項所述之系統，其中，該多層感知係使用該語音音段之驗證用特徵參數向量序列作為輸入，以一前饋動作以求出序列中每一驗證用特徵參數向量的驗證結果，並取該等驗證結果的平均值為該語音音段  
10 之驗證分數。

14. 如申請專利範圍第12項所述之系統，其中，該多層感知係使用一錯誤回傳學習演算法來縮小由該多層感知所輸出之驗證分數與一目標值間的均方根誤差，來訓練該多層感知。

15. 如申請專利範圍第14項所述之系統，訓練對應到一個驗證單元的多層感知係藉由輸入對應到該驗證單元的語音音段之驗證用特徵參數向量序列及非對應到該驗證單元的語音音段之驗證用特徵參數向量序列來訓練該多層感知，其中，訓練所使用的語音音段若對應到該驗證單元，  
20 則目標值為1，反之，則目標值為0。

16. 如申請專利範圍第10項所述之系統，其中，該驗證分數合併模組係取所有語音音段之驗證分數的平均值為該詞語驗證分數。

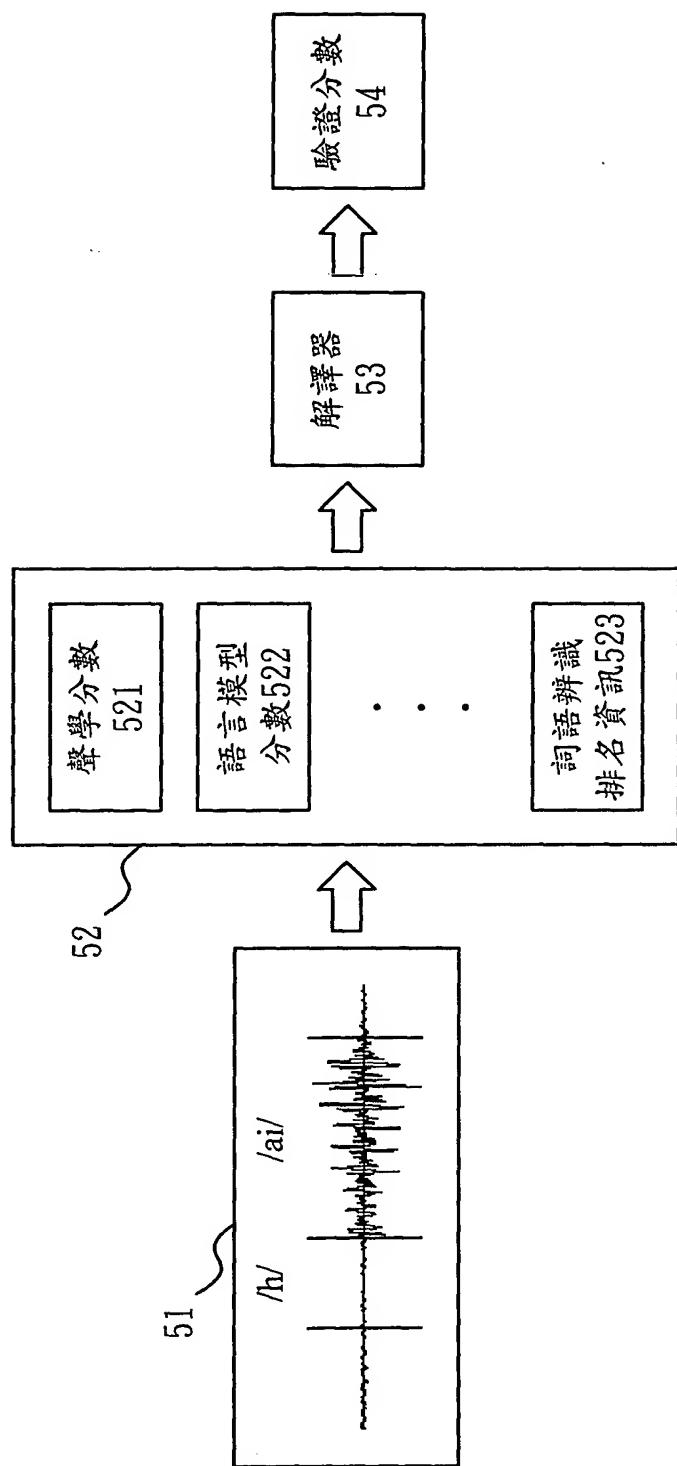


17. 如申請專利範圍第10項所述之系統，其中，該輸入語音信號係為一受噪音干擾之語音信號。

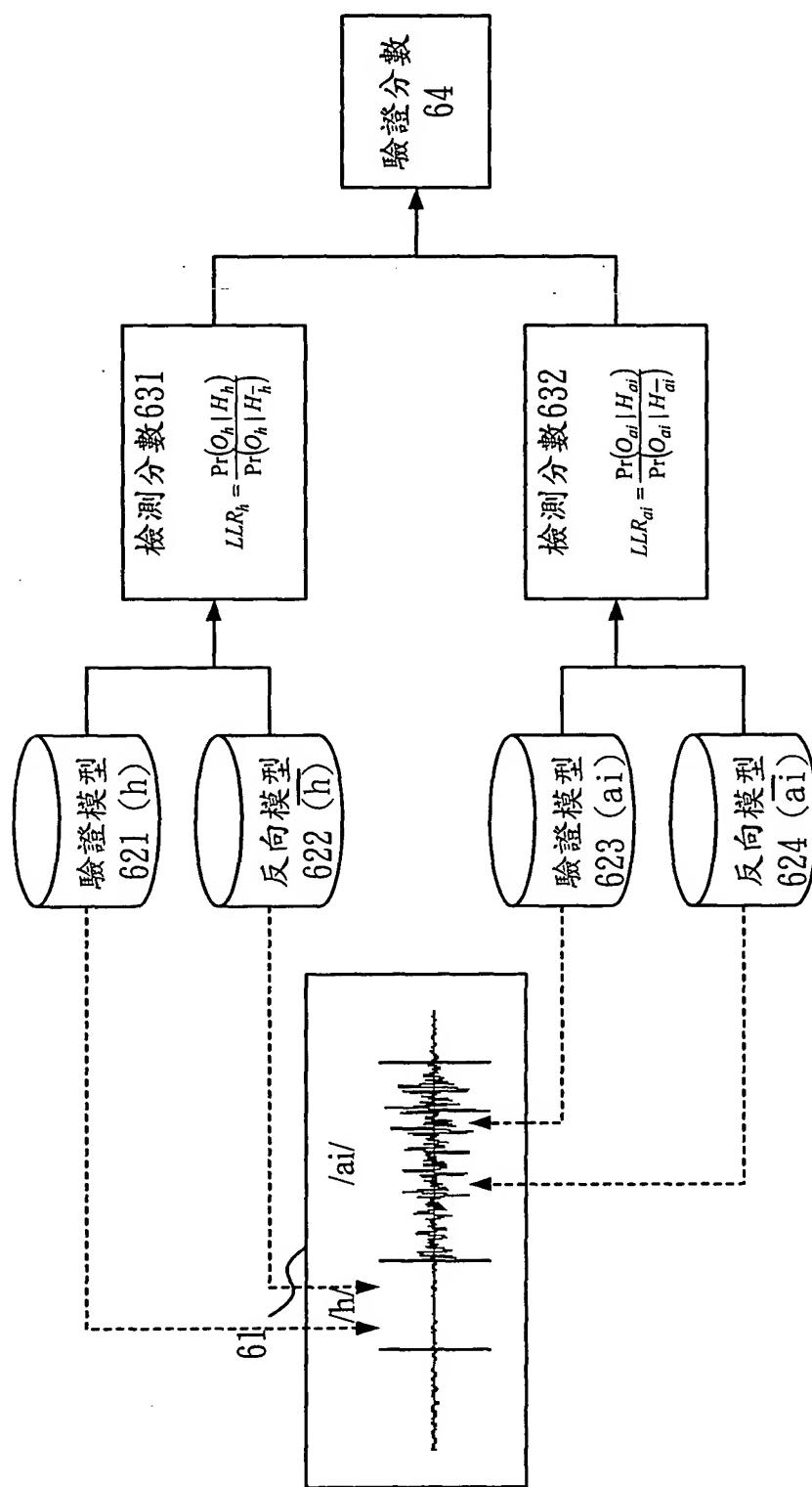
18. 如申請專利範圍第15項所述之系統，其中，該些為訓練所使用的語音音段係為受噪音干擾之語音音段。

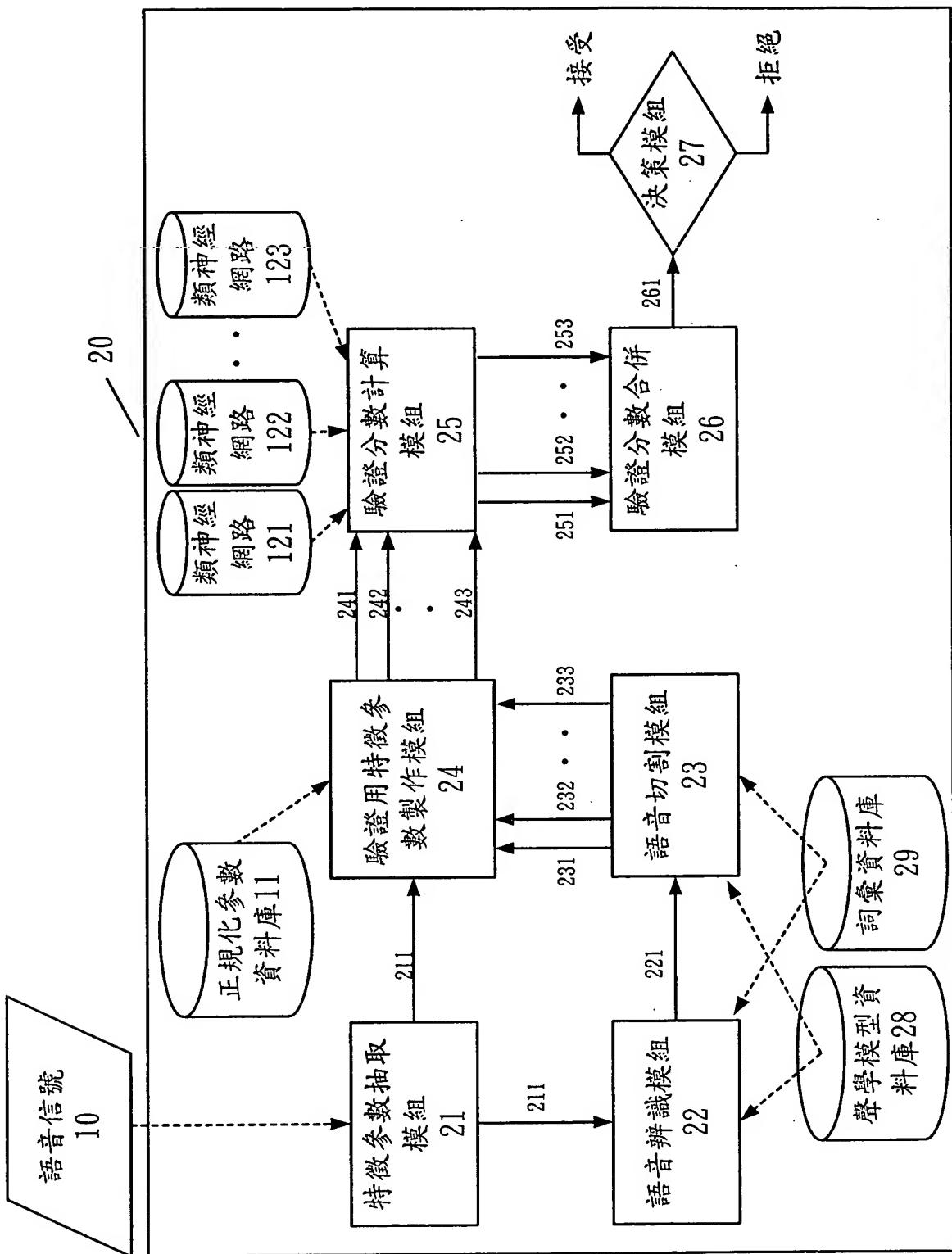


第1圖

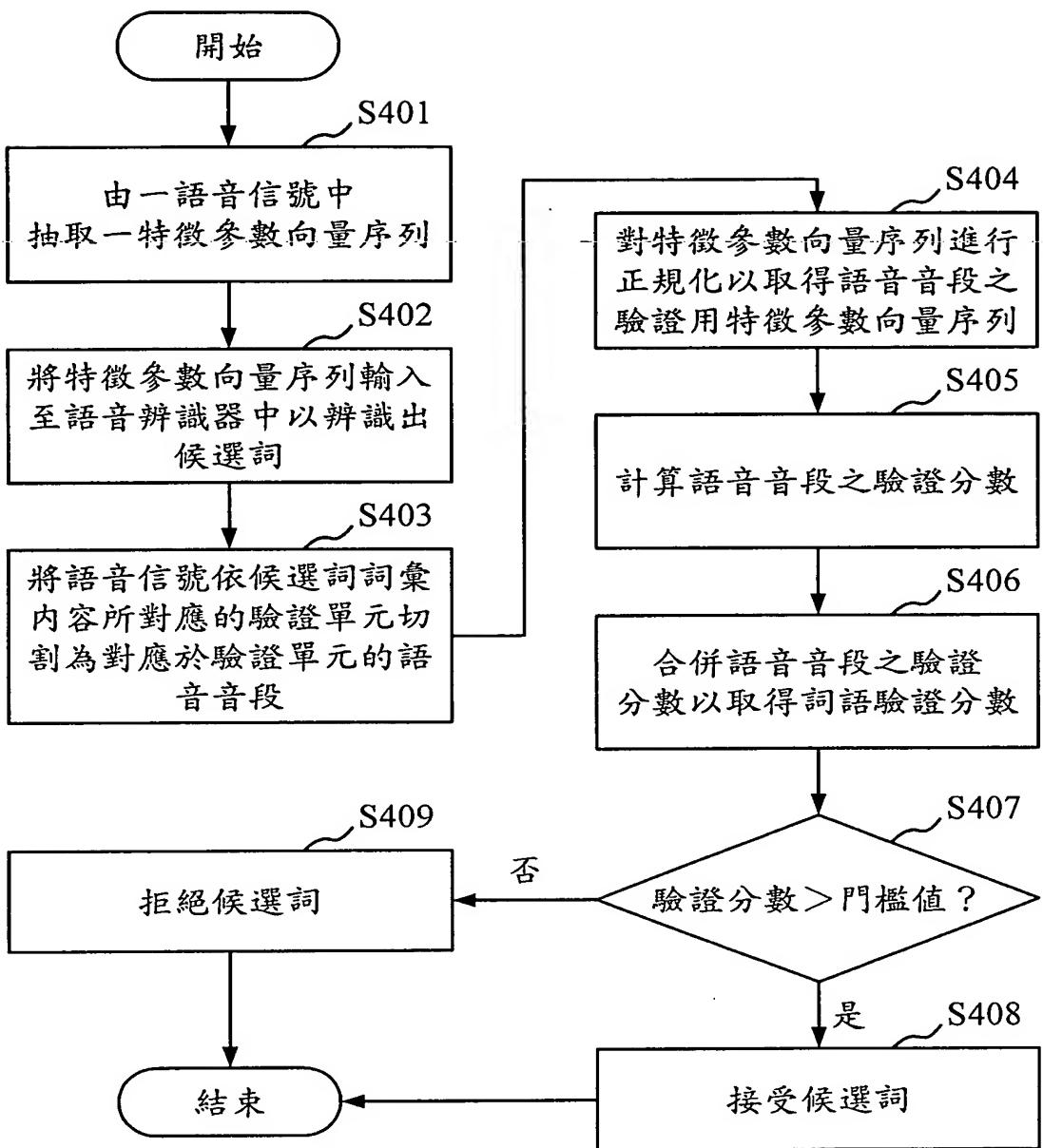


第2圖

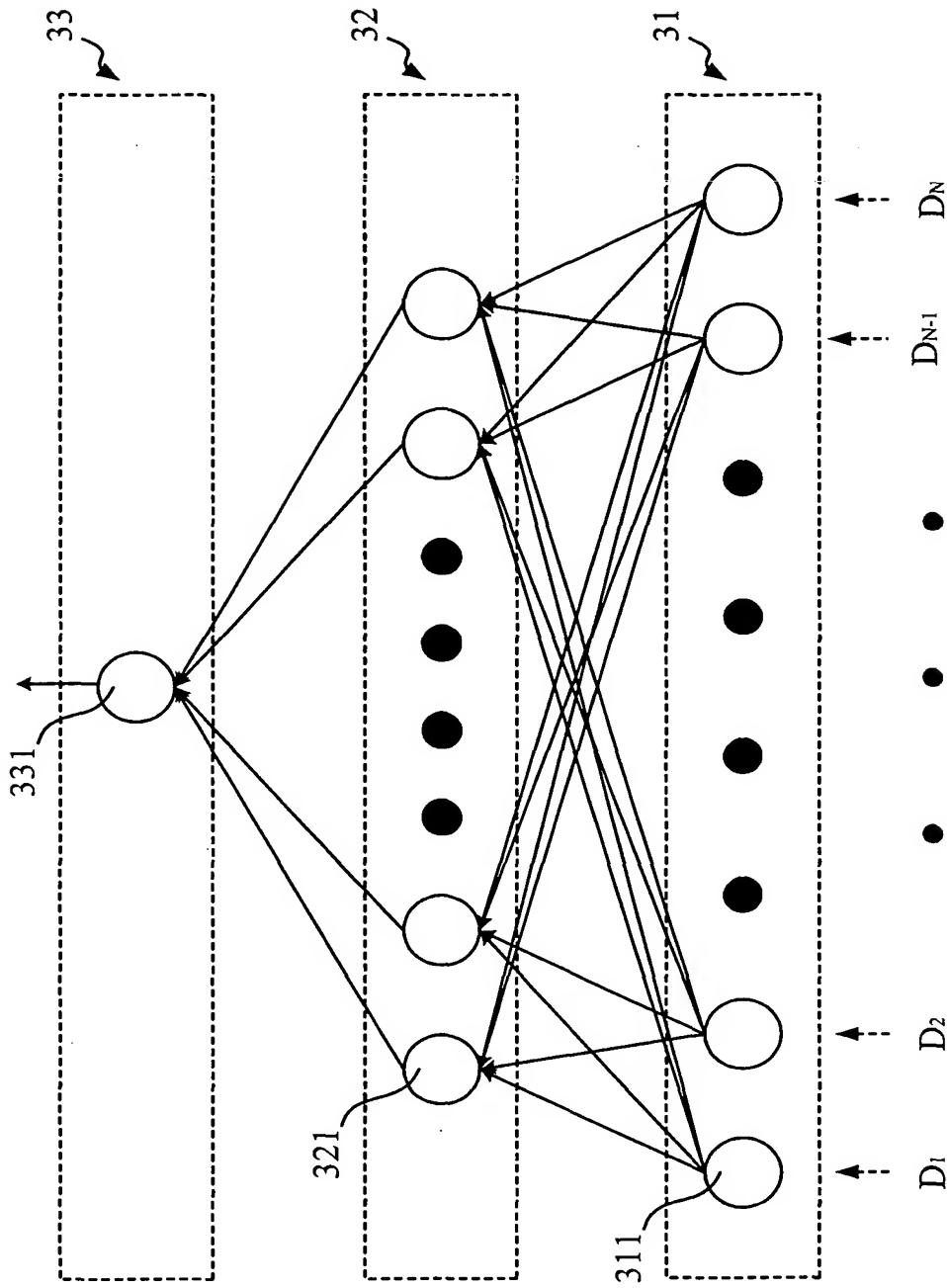




第3圖



第4圖



第5圖